

Score-based generative modeling for de novo protein design

Year: 2023 | Citations: 132 | Authors: Jin Sub Lee, Jisun Kim, Philip M. Kim

Abstract

The generation of de novo protein structures with predefined functions and properties remains a challenging problem in protein design. Diffusion models, also known as score-based generative models (SGMs), have recently exhibited astounding empirical performance in image synthesis. Here we use image-based representations of protein structure to develop ProteinSGM, a score-based generative model that produces realistic de novo proteins. Through unconditional generation, we show that ProteinSGM can generate native-like protein structures, surpassing the performance of previously reported generative models. We experimentally validate some de novo designs and observe secondary structure compositions consistent with generated backbones. Finally, we apply conditional generation to de novo protein design by formulating it as an image inpainting problem, allowing precise and modular design of protein structure. This study proposes a diffusion model, ProteinSGM, for the design of novel protein folds. The designed proteins are diverse, experimentally stable and structurally consistent with predicted models